AL/CF-TR-1995-0156

ARMSTRONG LABORATORY

# WORKLOAD RELATED CHANGES IN EYE, CARDIAC, RESPIRATORY AND BRAIN ACTIVITY DURING SIMULATED AIR TRAFFIC CONTROL (U)


Glenn F. Wilson

CREW SYSTEMS DIRECTORATE
HUMAN ENGINEERING DIVISION
WRIGHT-PATTERSON AFB OH 45433-7022

Carolyne R. Swain

LOGICON TECHNICAL SERVICES, INC.
P.O. BOX 317258
DAYTON OH 45431-7258

Jeffrey B. Brookings

WITTENBERG UNIVERSITY
DEPT OF PSYCHOLOGY
SPRINGFIELD OH 45503





AUGUST 1995

INTERIM REPORT FOR THE PERIOD FEBRUARY 1994 TO AUGUST 1995





AIR FORCE MATERIEL COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433-6573

## NOTICES

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Please do not request copies of this report from the Armstrong Laboratory. Additional copies may be purchased from:

National Technical Information Service
5285 Port Royal Road
Springfield, Virginia 22161

Federal Government agencies and their contractors registered with the Defense Technical Information Center should direct requests for copies of this report to:

Defense Technical Information Center
8725 John J. Kingman Road, Suite 0944
Ft. Belvoir, Virginia 22060-6218

## TECHNICAL REVIEW AND APPROVAL

AL/CF-TR-1995-0156

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

The voluntary informed consent of the subjects used in this research was obtained as required by Air Force Instruction 40-402.

This technical report has been reviewed and is approved for publication.

**FOR THE COMMANDER**

Kenneth R Boff

**KENNETH R. BOFF,** Chief
Human Engineering Division
Armstrong Laboratory

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



**1. AGENCY USE ONLY** *(Leave blank)*

**2. REPORT DATE**
August 1995

**3. REPORT TYPE AND DATES COVERED**
Interim, Feb 94 to Aug 95

**4. TITLE AND SUBTITLE**

Workload Related Changes In Eye, Cardiac, Respiratory and Brain Activity During Simulated Air Traffic Control

**5. FUNDING NUMBERS**
C F41624-94-D-6000
PE 62202F
PR 7184
TA 14
WU 25

**6. AUTHOR(S)**

Glenn F. Wilson, Carolyne R. Swain *, Jeffrey B. Brookings **

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
Armstrong Laboratory, Crew Systems Directorate,
Human Engineering Division
Human Systems Center
Air Force Materiel Command
Wright-Patterson AFB OH 45433-7022

**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**

AL/CF-TR-1995-0156

**11. SUPPLEMENTARY NOTES**

* Logicon Technical Services, Inc., P.O. Box 317258, Dayton OH 45431-7258
** Wittenberg University, Department of Psychology, Springfield OH 45503

**12a. DISTRIBUTION/AVAILABILITY STATEMENT**

Approved for public release;
distribution is unlimited.

**12b. DISTRIBUTION CODE**

**13. ABSTRACT** *(Maximum 200 words)*

In this investigation, eight Air Force air traffic controllers (ATCs) performed three scenarios on TRACON, a computer-based ATC simulation. Two scenarios, each with three levels of difficulty, varied either traffic volume by manipulating the number of aircraft to be handled or varied traffic complexity by manipulating arriving to departing flight ratios, pilot skill and the mixture of aircraft types. The third scenario, referred to as the overload condition, required that controllers handle an extremely high traffic volume in a very limited amount of time. The effects of the manipulations on controller workload were assessed using performance-based, subjective (TLX), and physiological (heart, eye, respiration and EEG) measures. Significant main effects of difficulty were found for TRACON performance, TLX ratings, eye blink rates, respiration rates and EEG measures. Only the EEG data were associated with main effects of traffic pattern. The results provide support for the differential sensitivity of a variety of workload measures in complex tasks, underscore the importance of traffic complexity in ATC workload, and support the utility of TRACON as a tool for studies of ATC workload.

**14. SUBJECT TERMS**
Psychophysiology, Performance Measures, Air Traffic Control, Cognitive Workload, Electroencephalography

**15. NUMBER OF PAGES**
40

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| UNCLASSIFIED | UNCLASSIFIED | UNCLASSIFIED | UNLIMITED |

NSN 7540-01-280-5500



Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

THIS PAGE LEFT BLANK INTENTIONALLY

TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

# PREFACE

This report was prepared in the Human Engineering Division, Crew Systems Directorate, of the Armstrong Laboratory (AL) under Project 7184, Task 1425, "Physiological Workload Assessment." We would like to thank the Wright Patterson Air Force Base air traffic controllers who participated in this study. We also acknowledge the help of Barbara Palmer and George Reis with data collection and Chuck Goodyear for his statistical expertise. When this research was conducted, Dr. Brookings was serving at Armstrong Laboratory on a National Research Council Research Associateship. This project was supported in part by funds from the FAA Technical Center with Dr. Ed Buckley and Dr. Earl Stein serving as the program managers.

# INTRODUCTION

It has long been recognized that air traffic control (ATC) is a complex and demanding job (see Noland, 1990, for a discussion of ATC tasks and responsibilities). Over the past decade, amid growing concerns that increased traffic in the national airspace system threatens to overwhelm ATCs and compromise the safety of air travel (Danaher, 1985), more attention has been focused on ATC workload. For example, the Federal Aviation Administration (FAA) developed the National Airspace System plan that proposed, among other things, to increase the automation of ATC (FAA, 1985). It was anticipated that automation would decrease ATC workload, so that projected increases in the demand for air traffic services could be met safely and efficiently.

The notion that decreased workload should result in improved ATC service is intuitively reasonable, but Hopkin (1989) noted that the salutary effects of workload reduction "...seem to be assumed rather than proved" (p. 103). Indeed, studies of workload and ATC operational errors have yielded mixed results. Morrison and Wright (1989) analyzed reports from NASA's Aviation Safety Reporting System, and found that controller errors (e.g., monitoring failures, improperly executed handoffs, wrong heading or altitude assignments) were associated most frequently with increases in workload factors such as traffic volume and frequency congestion. On the other hand, Redding's (1992) analysis of FAA data for 1989 indicated that ATC operational errors occurred more frequently under moderate--rather than high--workload conditions (see also Stager, Hameluck, & Jubis, 1989).

Statistical analyses of incident data are useful for characterizing overall ATC system performance and for generating workload evaluation hypotheses. But, the need for controlled studies of ATC workload is apparent when consideration is given to a) the reliance of such analyses on accurate and complete incident reports, and b) in retrospective analyses controller workload must be inferred from traffic data, rather than assessed directly. To date, much of the research employing physiological measures to study ATC has been focused upon issues related to the effects of long term stress on controller health. While this is quite

important, the focus of this study is on the effects of the momentary changes in workload experienced by controllers as they perform their duties.

This approach is more directly related to the concerns stated above about the effects of increased air traffic as well as the effects of automation upon the workload of the individual controller and the application of these metrics in comparable situations is not novel. A variety of psychophysiological metrics have been used in a number of real world environments to monitor operator workload (see Wilson & Eggemeier, 1991, for a review). These have included, but not been limited to, measures of eye (EOG), heart (ECG), respiratory and brain (EEG) electrical activity. For a review of various techniques and their application to the assessment of mental workload, see Ullsperger, 1993. For a review of psychophysiological methods, see Caldwell et al, 1994.

Heart rate has been reported to vary as a function of the mental load imposed by operator's task. Wilson (1993) reported that increased heart rate was associated with more difficult aspects of fighter aircraft air-to-ground missions and further that the heart rate of the pilot increased more than that of the accompanying weapons systems officer (WSO) except when the WSO was in control of the aircraft. In the realm of aircraft systems evaluations, heart rate measures have been used in the certification of civilian airliners (Roscoe, 1987; Speyer et al., 1988). The beat-to-beat variability of the heart rhythm has also been used as a measure of mental effort and has been reported to decrease with increases in mental demand in environments ranging from simulated flight (Itoh et al., 1989) to car driving (Egelund, 1982). However, not all studies have reported heart rate variability to be sensitive to levels of mental demand (Wilson, 1992).

Measures of respiration rate and amplitude have also been applied to the work environment and have been reported to change in conjunction with mental demand (for reviews see Caldwell, et al., 1994 or Wientjes, 1992). In general, it has been demonstrated that respiration rates increase and the depth of respiration decreases during periods of increased cognitive demand but these responses are susceptible to adaptation and training so

that highly trained individuals may not readily evidence changes in response to subtle or short term variations in cognitive workload (see Wilson, 1992).

Several investigators have reported decreased eyeblink rates in simulated flight situations with high visual demands (Stern & Dunham, 1990; Stern & Skelly, 1984) and Wilson et al. (1987) and Skelly et al. (1987) both reported decreased blink rates and closure duration during higher workload segments in actual and simulated flight. Laboratory results, however, have not been as consistent. For example, Bauer et al. (1985) and Casali & Weirville (1983) failed to report a relationship between task demands and blink rate. In addition, reports have linked an increase in slow eye movements (saccades) and an increase in blink rate and closure duration to fatigue and time on task (Caldwell et al., 1994).

Measures of the brain's electrical activity are seen by many to be direct ways of determining the cognitive demands placed upon an operator. The brain is responsible for processing information, making decisions and initiating actions on the external environment. Thus it is reasonable to expect that alterations in the observed electrical activity generated during task performance reflect corresponding alterations in mental task demands. Without question, ATC is a complex and demanding task, no doubt involving numerous brain areas. Consequently, recording from multiple sites over the brain is warranted. Because the psychophysiological measures are continuously present and their recording is not intrusive to the operator's job performance, it was decided to use them as measures of controller workload as suggested by Kalsbeek (1971).

Many laboratory and simulator studies have used untrained subjects or college students who were trained on part-task ATC simulators. This highlights a recurrent problem when investigating controller workload. It is both expensive and time consuming to select and train subjects in order to establish a test group equivalent to ATC operators. To help alleviate the problems inherent in using partially trained subjects, Air Force ATCs were recruited to participate in this study. These controllers had already been selected and trained and thus possessed the strategies and skills commonly used by controllers in the performance

of their jobs. The purpose of the current investigation was to assess workload related changes in psychophysiological responses associated with variations in the difficulty of TRACON (Terminal Radar Approach Control; Wesson International), a simulated ATC task. Multiple measures--performance, subjective, and physiological--were collected to provide a comprehensive perspective on controller workload.

## METHODS

### Subjects

Eight Air Force ATCs (7 male, 1 female) volunteered and were paid for their participation. All subjects were right handed, ranged from 21 to 29 years of age, and reported ATC experience of 2.5 to 7.5 years.

### Apparatus

The TRACON scenarios were administered on a Unisys 386/25 Mhz processor and ViewSonic 7 color VGA monitor (17", 640H X 480V resolution graphics). Subjects entered their commands via keyboard and a two-button mouse. Pilots' voices, which were generated using a SoundBlaster interface card, were presented over a speaker.

Electrophysiological recording of peripheral measures was accomplished using the Psychophysiological Assessment Test System (Wilson & Oliver, 1991). EOG was monitored for vertical blink and horizontal saccade activity via Ag/AgCl electrodes positioned above and below the left eye and at the outer canthus of both eyes. EOG signals were filtered at 0.1 - 30 Hz and amplified by 5000 using Grass P511 amplifiers. Ag/AgCl electrodes were positioned on the sternum and fifth intercostal space on the left side of the body to record heart activity. Cardiac signals were amplified at 2000 and the data were filtered at 10 - 100 Hz. The ground electrode was positioned on the right side of the ribcage in the fifth intercostal space. Respiration was monitored using a Respitrace system with elastic transducer bands on the chest and abdomen which were summed for output. Prior to data collection, respiration amplitude was calibrated for each subject. Nineteen channels of brain wave data were recorded at sites positioned according to the International 10-20 electrode system (Jasper 1958), using an ElectroCap and Biologic Brain Atlas III. Linked mastoids

served as reference. The amplifier gain was 30,000 with bandpass filters at 0.1 and 30 Hz. Electrode impedances were maintained at < 5 Kohms.

## Task

The simulated ATC task was TRACON for Windows (Version 1.03, Wesson International). The TRACON display consisted of: a) a color radarscope depicting Los Angeles International and four surrounding airports; b) flight strips listing "active" and "pending" aircraft; c) a communications box that provided an ongoing visual record of controller commands and pilot responses; and d) the controller's score for the current scenario. All information displayed on the radarscope (e.g., airways, VOR radio beacons, sector boundaries) represented accurately the information contained in government airspace charts.

The subjects' task was to handle a series of aircraft requesting ATC services by issuing commands (e.g., turn, descend) that enabled the safe and efficient execution of their respective flight plans. The flights, representing a variety of aircraft types flying under either instrument or visual flight rules, included arrivals, departures, and overflights. Subjects received points for successfully handling aircraft, minus any points deducted for operational errors (separation conflicts, hand-off errors, missed approaches).

In an initial briefing, controllers were introduced to the TRACON simulation and were given instructions on how to issue commands using the keyboard and mouse. During off-duty hours, the controllers then completed a series of eight practice simulations arranged in order of increasing difficulty. Subjects were required to complete each simulation with no crashes, separation conflicts, or hand-off errors before proceeding to a more difficult simulation. The controllers maintained personal performance records for each practice trial, and the mean time required to complete the eight simulations was approximately six hours. After all practice simulations were completed successfully, controllers were scheduled for their experimental session.

In the experimental sessions, each subject completed three scenarios. The rationale for the workload manipulations underlying these scenarios was derived from a review of

ATC literature and conversations with civilian and Air Force ATCs. One scenario consisted of 36 aircraft presented over a 45-minute period. It included 15 minutes each of low, medium and high workload segments defined by total traffic volume (6, 12, and 18 aircraft, respectively). All other factors, such as aircraft type and the ratio of arrivals to departures and overflights, were held constant across the segments. The second scenario was also 45 minutes long and consisted of 36 aircraft equally divided among three workload segments, but the number of aircraft (12) was held constant. Instead, the segments varied in degree of traffic complexity, operationalized as a function of: a) the ratio of arriving flights to departures and overflights; b) the probability that pilots either did not hear or failed to execute properly the controller's commands; and c) the heterogeneity of aircraft type. In both scenarios, the workload segment transitions were arranged so that subjects had sufficient time, approximately one minute, at the conclusion of a segment to complete subjective workload ratings. The third scenario, an overload condition, presented controllers with 15 aircraft in 5 minutes. The purpose of this scenario was to overwhelm the controllers or, in ATC parlance, make them "lose the picture" with particular attention to physiological activity accompanying this phenomenon.

## Dependent Measures

The principal performance measure was the ratio of TRACON points earned to the total points possible for that segment. The subjective workload measures were the NASA TLX scores (Hart & Staveland, 1988), based on ratings collected at the end of each segment. The TLX contains six subscales that are scored from a low of 0 to a high of 100: mental demand, physical demand, temporal demand, performance, effort and frustration. A composite score is also derived based upon combining the scores from all of the subscales. In order to keep the scenarios realistic, they were designed so that the workload in each condition builds to maximum at approximately the midpoint of the scenario and then tapers as the controller successfully handles aircraft entering his/her airspace. Consequently, the peripheral physiological measures were calculated from 5-minute intervals at the mid-point of the workload periods within each segment and included: a) heart rate, heart interbeat interval (IBI), and heart rate variability; b) eye blink rate, interblink interval, and blink amplitude; c) saccade rate, amplitude and mean duration; and d) respiration rate and

amplitude. FFTs were calculated based upon two consecutive one minute periods of EEG data following the peak workload of each segment. Before the FFTs were computed, EEG data were corrected for eye movements using a modified version of the procedure developed by Gratton, Coles & Donchin (1983) and portions of EEG data which contained other artifacts were discarded. FFT data were then collapsed into five bands for further analysis. The bands were: delta (1.1-3.9 Hz ), theta (4.3-7.8 Hz ), alpha (8.3-11.9 Hz), beta 1 (12.3-15.8 Hz ), and beta 2 (16.2-24.9 Hz ). To accommodate for intersubject variations in absolute EEG power, spectral band values were expressed as the percentage of the total power between 1.1 to 24.9 Hz. SAS was used to generate statistical comparisons. Repeated measures analyses of variance (ANOVAs), with the Geiser-Greenhouse correction, were computed to assess main effects. Pairwise contrasts (t tests) were used to test differences between individual conditions. Those reported are significant at $p < .05$.

## Procedure

Each controller's experimental session lasted approximately four hours. This included: a) the time required to instrument the subjects (i.e., attach electrodes, affix the respiration bands and EEG cap); b) a three-minute resting period to collect physiological baseline data; c) a five-minute warm-up simulation (four aircraft) to familiarize controllers with the laboratory computer; d) the two long scenarios; e) the overload scenario; and f) 15-minute breaks between each of the three principal scenarios.

The order in which controllers completed the volume and complexity scenarios was counterbalanced and the order of the three workload segments within each scenario was also counterbalanced as completely as possible given the relatively small sample size. Due to the potential for detrimental effects, the overload scenario was always completed last. Experimental sessions were conducted in an electrically shielded and sound attenuated chamber. Each session was videotaped for off-line evaluation of subject's responses and performance. The experimenter sat behind and to the right of the subject, and administered the TLX after each workload segment of the long scenarios and at the conclusion of the overload scenario.

## RESULTS

For the performance and subjective data, ANOVAs indicated: 1) significant main effects of task difficulty (low, medium, high) on TRACON performance, all six of the TLX subscales, and the TLX composite score (see Table 1 for a summary); 2) no main effects for traffic (volume vs complexity); and 3) no significant difficulty x traffic interactions. TRACON and TLX composite score data are presented in Figures 1 and 2, respectively. Performance scores in the high volume and high complexity conditions were significantly lower than the low complexity condition. Further, the overload condition produced lower performance scores than the low and medium workload conditions of either volume or complexity scenarios. The TLX composite score results mirrored the performance data with the addition of high complexity condition being rated as more difficult than the medium complexity condition and a significant difference between low and medium volume conditions.

| Scale | F-Value | Significance |
|---|---|---|
| Mental Demand | 10.01 | $p < .0020$ |
| Physical Demand | 10.02 | $p < .0020$ |
| Temporal Demand | 10.31 | $p < .0018$ |
| Performance | 12.98 | $p < .0006$ |
| Effort | 21.03 | $p < .0001$ |
| Frustration | 11.44 | $p < .0011$ |
| TLX Composite | 25.65 | $p < .0001$ |
| TRACON Score | 4.64 | $p < .0285$ |

**Table 1.** A significant effect of task difficulty (low, medium, high) was evidenced on all TLX subscales, the TLX composite score and the proportion of TRACON points earned.

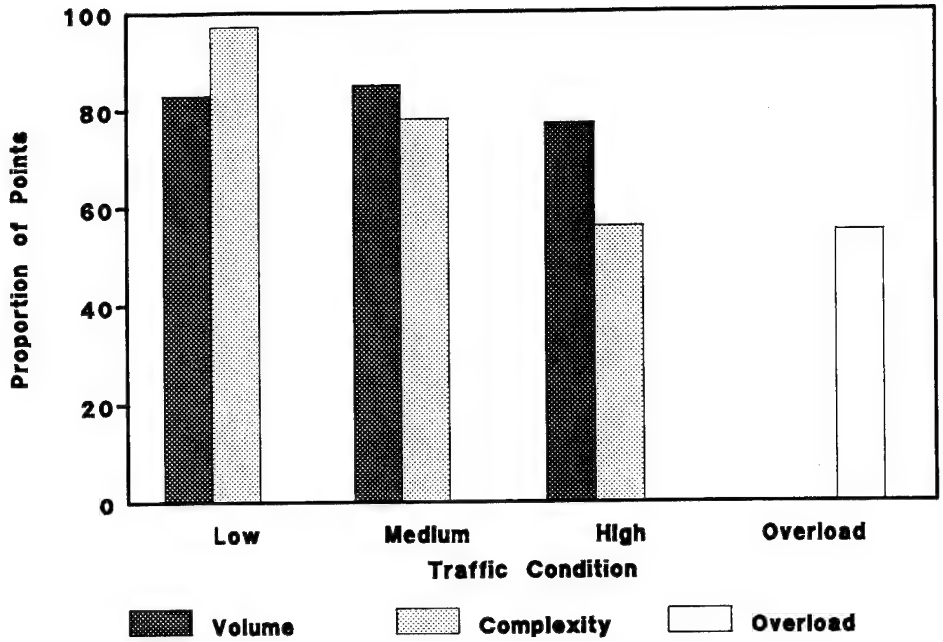



**Figure 1.** Performance data expressed as a proportion of total possible TRACON points for each level of difficulty and traffic pattern.

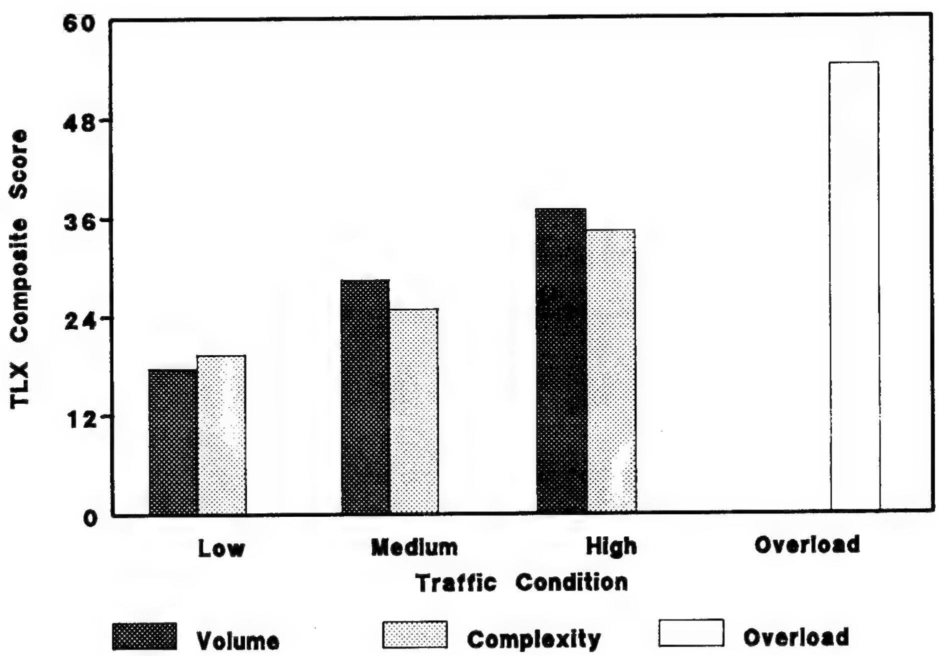



**Figure 2.** Composite subjective ratings for each level of task difficulty and all traffic scenarios.

Repeated measures ANOVAs of the peripheral psychophysiological measures indicated: 1) significant main effects for workload on eye blink rate but not interval or amplitude, and a significant main effect for workload on both respiration rate and amplitude; 2) no main effects for traffic; 3) a significant workload x traffic interaction for respiration amplitude; and 4) no significant main effects or interactions for any of the heart or saccade measures.

Blink rate decreased significantly as the TRACON task became more difficult ($F=9.37$, $p<.0.01$). However, there was no significant difference between the two traffic manipulations ($F=2.75$, $P>0.05$). As can be seen in Figure 3, the blink rates declined with increasing TRACON difficulty. Post hoc tests demonstrated that the blink rates in the low level conditions were significantly higher than those observed during the high workload levels or during the overload condition. The medium volume condition blink rates were significantly lower than the low volume level. The medium complexity condition blink rates were significantly lower than the low volume condition and higher than the high complexity condition. The blink rates recorded from the overload condition were significantly lower than both low workload conditions and the medium complexity condition. There were no significant changes in blink amplitude or duration (Figure 4).

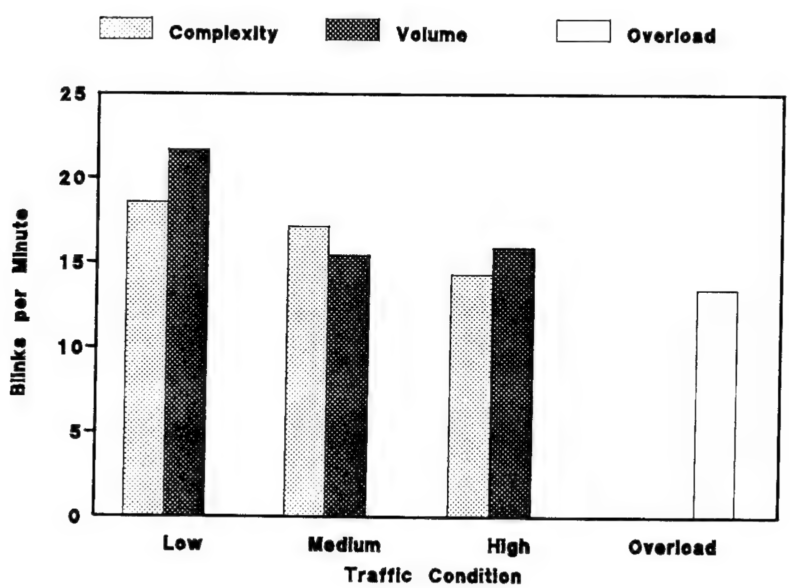



**Figure 3.** Average number of blinks per minute for each TRACON segment.

a.

b.

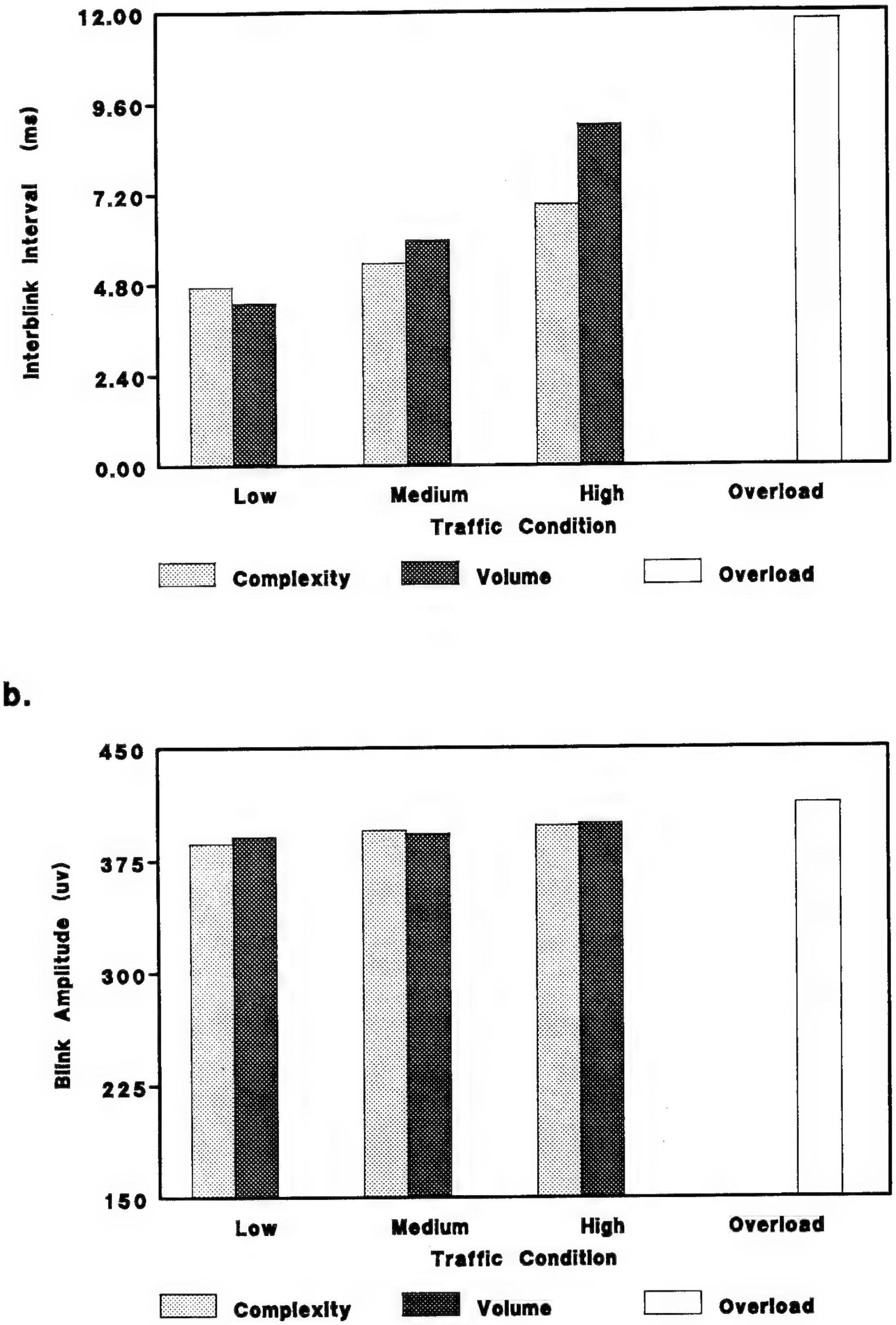



**Figure 4.** a) The average blink interval did not differ across segments. b) The average blink amplitude was nearly equivalent in all segments.

Although respiration amplitude was unaffected, respiration rate was higher as the TRACON difficulty increased ($F=17.88$, $p<0.0015$). However, respiration rates were not affected by the traffic manipulation ($F=1.84$, $p<0.22$). The mean respiration rates are depicted in Figure 5 and the mean respiration amplitudes are shown in Figure 6. The high complexity condition was associated with significantly higher respiration rates than both low conditions and the medium volume condition. Both medium conditions and the overload condition showed significantly higher rates compared to the low complexity condition while the overload respiration rates were also significantly higher than those of the low volume condition.

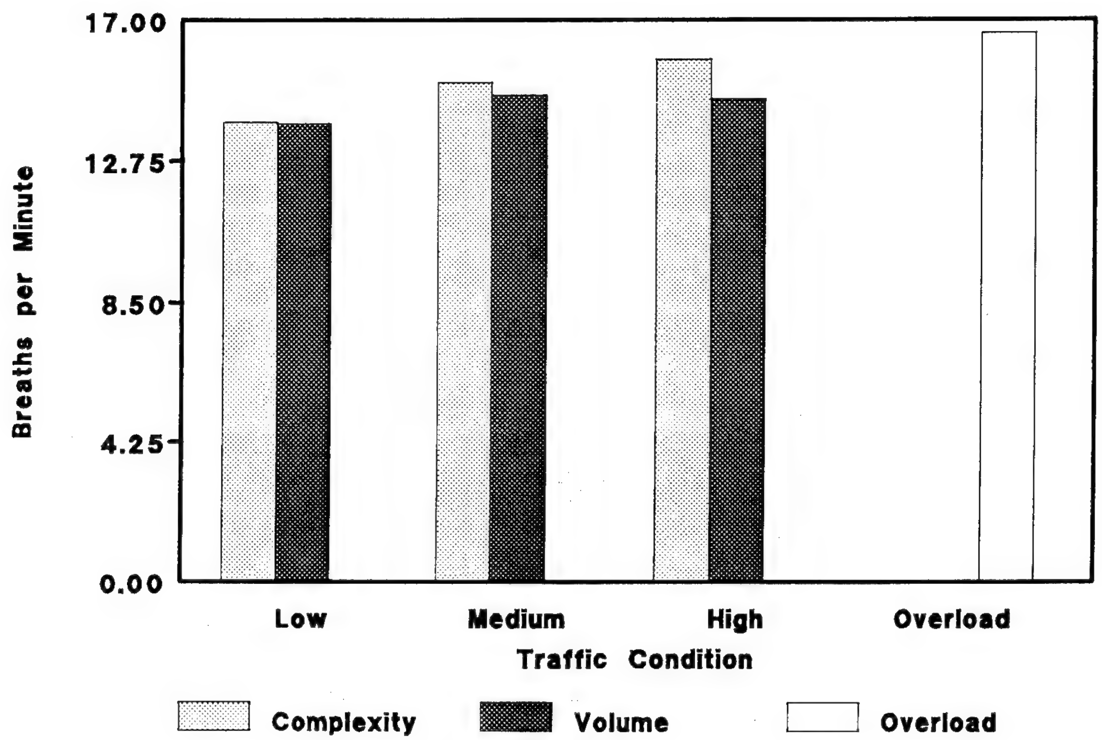



**Figure 5.** The average respiration rates for each scenario and level of difficulty.

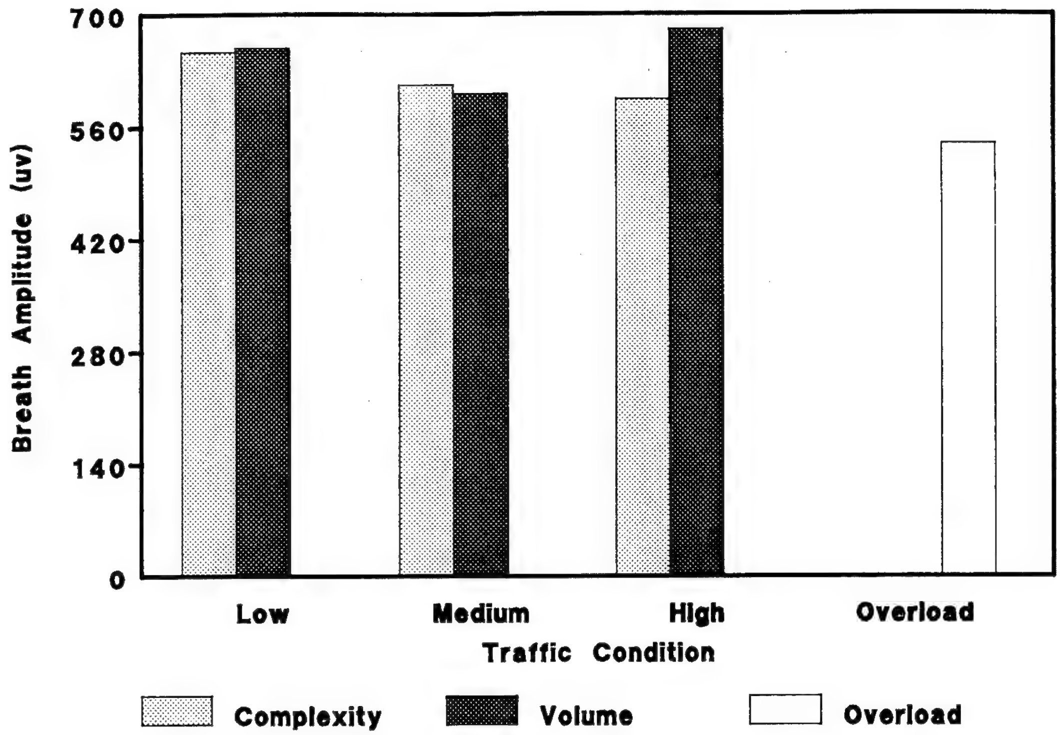
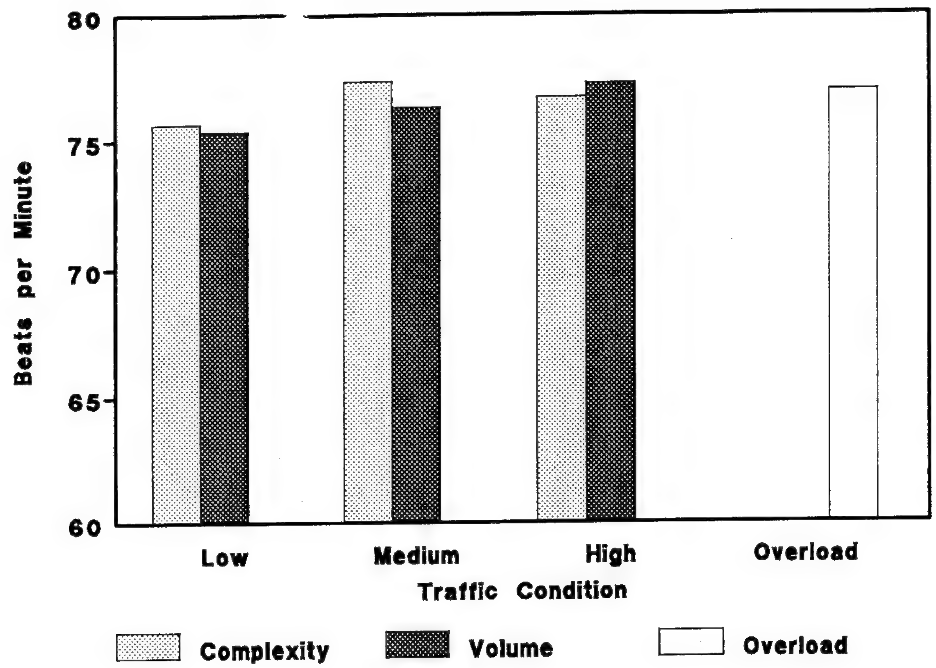



**Figure 6.** The average respiration amplitudes did not differ significantly between conditions.

Heart rate measures did not demonstrate significant differences in response to either difficulty or traffic manipulations (see Figure 7). All TRACON conditions were associated with heart rates higher than the resting baseline by 1.77 to 3.73 beats per minute. Heart rate variability in the 0.15 to 0.4 Hz, or respiration, band did approach significance as a result of the difficulty manipulation ($F=3.54$, $P<0.066$) but there were no task related differences in the 0.06 to 0.14 Hz blood pressure band (Figure 8). Measures of slow eye movement, or saccades (rate, amplitude and duration) also demonstrated no significant changes due to any experimental manipulations (Figure 9).

The FFTs for the two one-minute segments following the "peak" of each condition did not differ statistically and were pooled for the analyses. Figure 10 summarizes the results of the ANOVAs performed on the EEG data from the nineteen channels for the five bands.

a.




b.

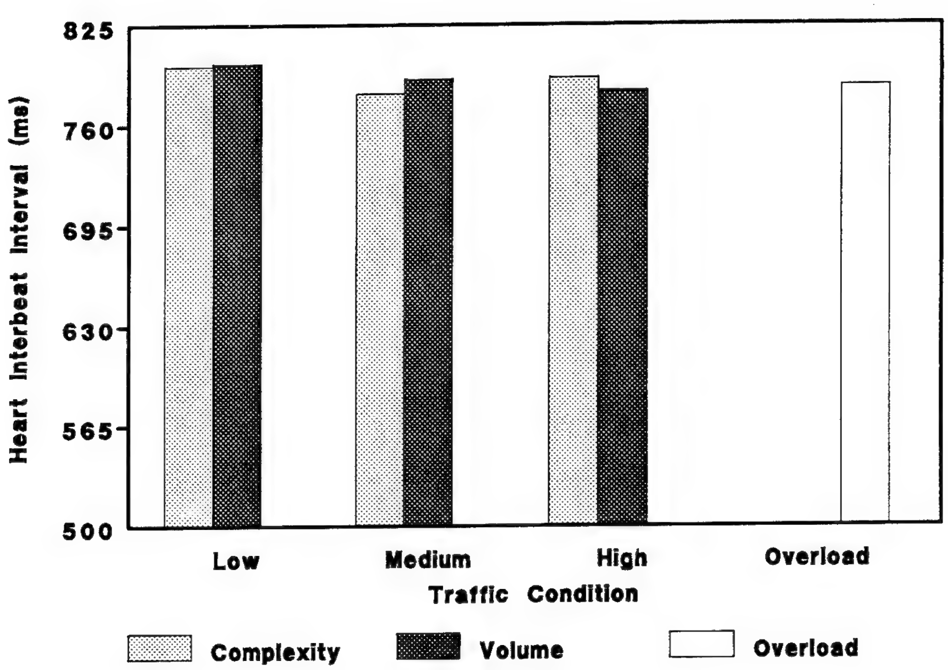
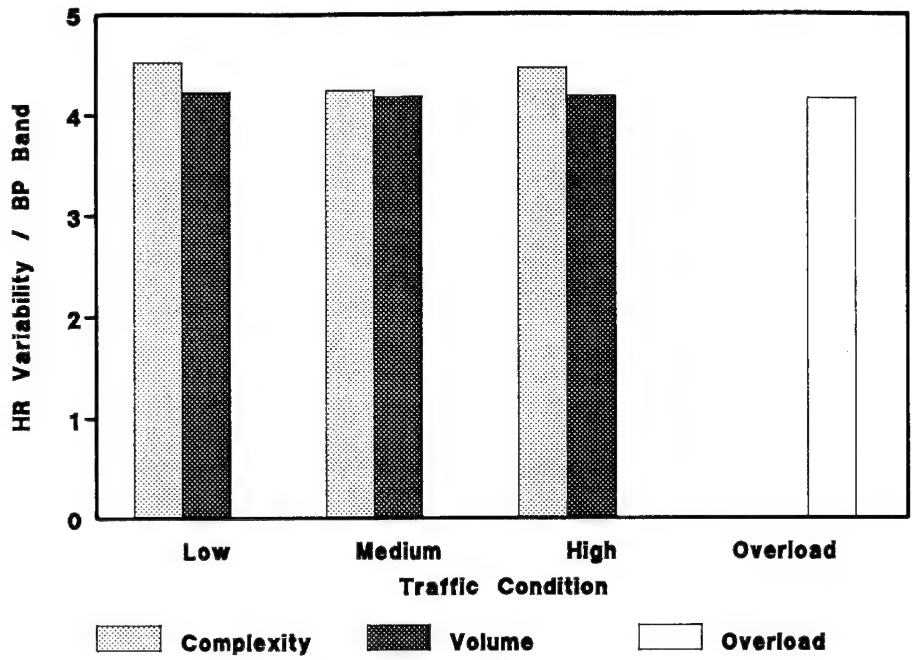



**Figure 7.** a) No differences were noted in the average heart rate across conditions. b) The average interbeat interval was nearly equivalent in all conditions.

a.




b.

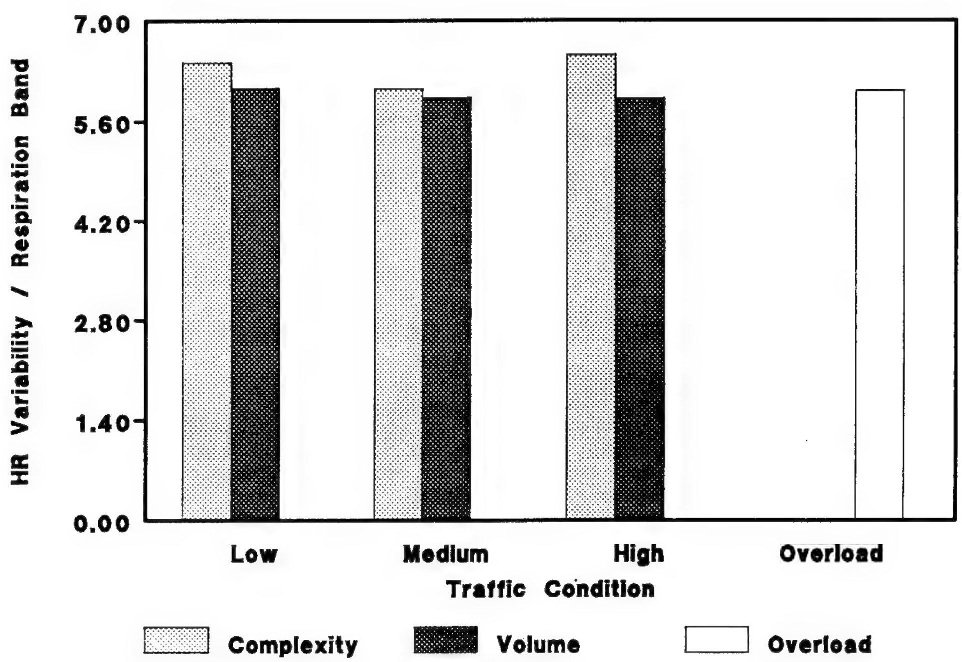



**Figure 8.** a) Heart rate variability in the .06 to .14 Hz band did not differ between conditions. b) Heart rate variability in the .15 to .40 Hz band did approach significance as a result of the manipulations in task difficulty.

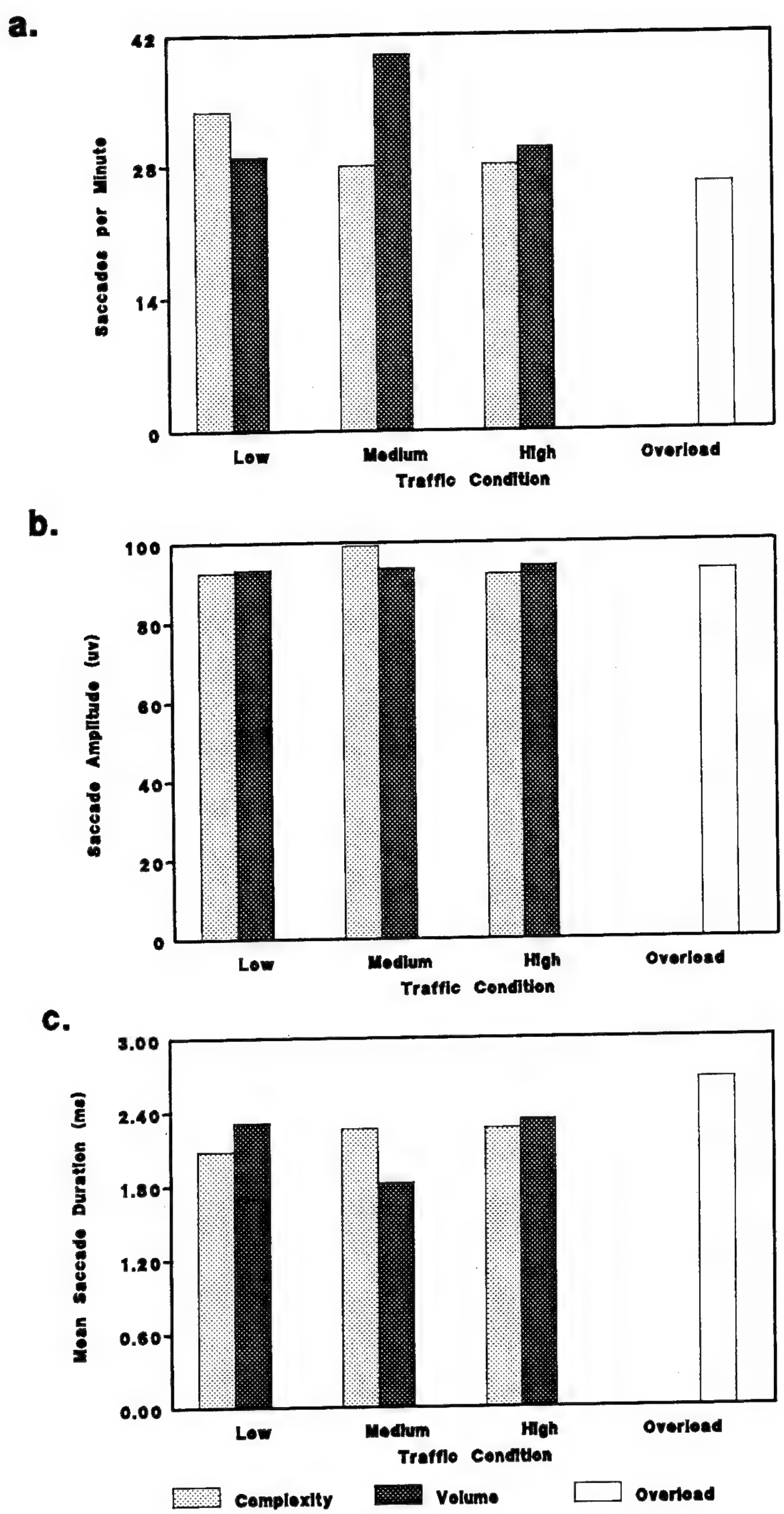



**Figure 9.** a) The average number of saccades per minute did not differ during conditions. b) Average saccade amplitude was unaffected by task manipulations. c) The duration of saccades did not change as a result of task requirements.

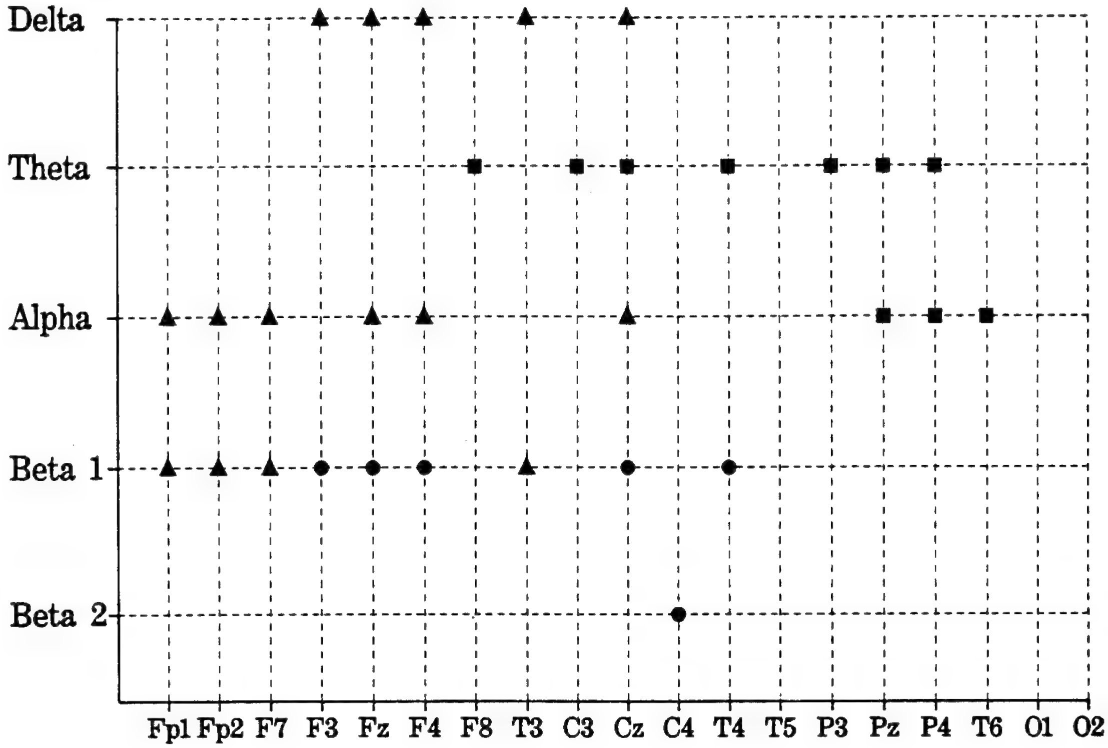



**Figure 10.** Significant F-values are indicated by a circle (traffic manipulation), box (difficulty manipulation), or triangle (interaction) for each combination of electrode and EEG band. Intersections without a symbol did not reach significance.

The delta band activity demonstrated significant differences due to the interaction between traffic and difficulty manipulations at electrode sites F3, Fz, F4 and T3. For the traffic complexity scenario, the highest percent delta power was found during the medium difficulty condition with lowered values associated with the high condition. The low difficulty condition evidenced the least delta power with significant differences between medium and low at Fz and F4. For the traffic volume scenario, the low difficulty segment was associated with higher power than medium and high difficulty segments which were approximately the same magnitude. Pairwise contrasts indicated low difficulty responses to be significantly enhanced compared to high difficulty responses only at electrode site T3. See Figure 11 for a summary.

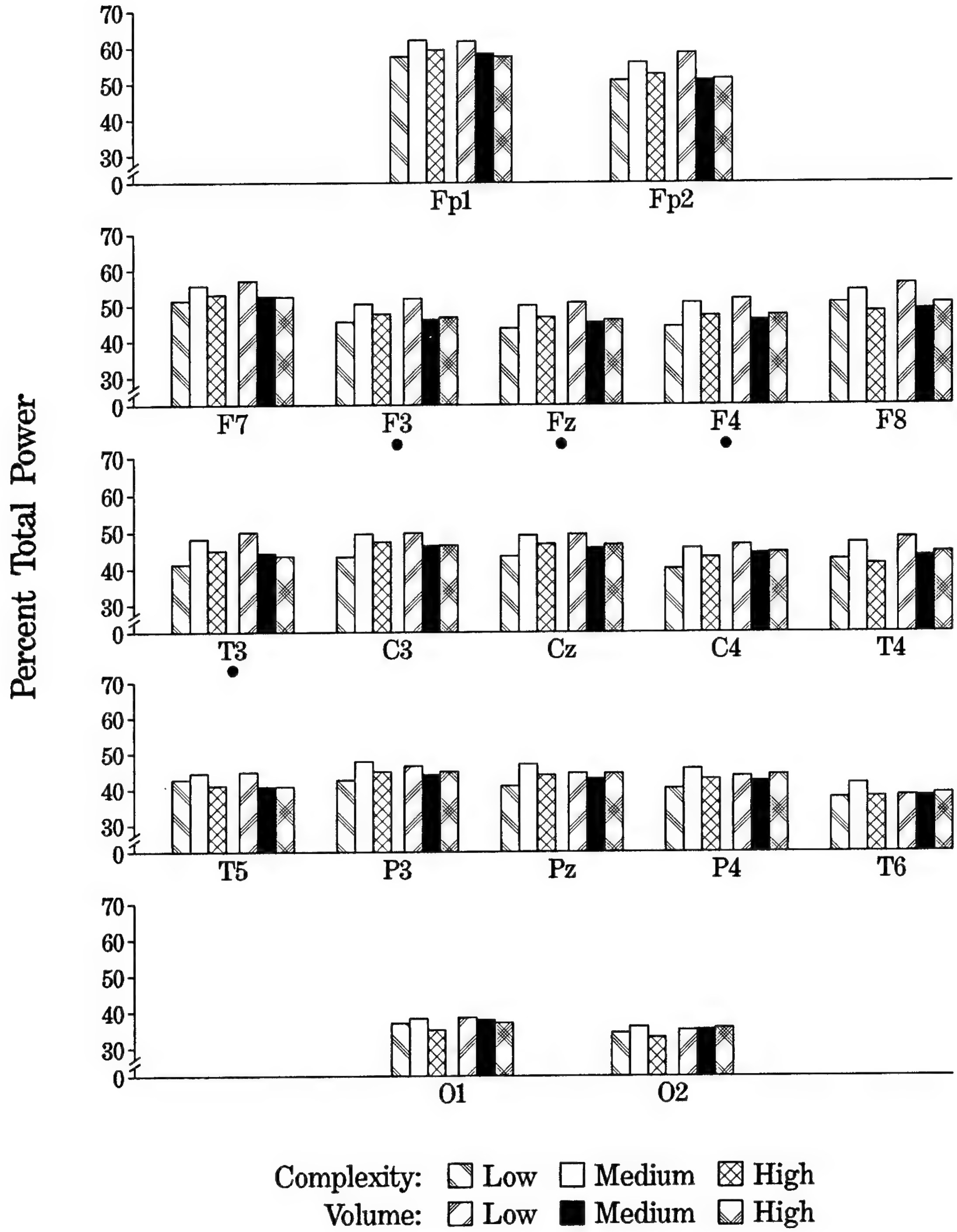



**Figure 11**. Percent total power from the delta band for each electrode at each level of difficulty during the volume and complexity scenarios. The filled circles below an electrode indicate a significant interaction at p <.05.

Figure 12 demonstrates that enhanced theta band activity was seen primarily at central and parietal sites with similar increases apparent at right frontal and right temporal recording sites. Increased mental workload in both of the traffic manipulations was associated with increased theta relative power. The high difficulty situations were always associated with significantly greater relative power than the low difficulty conditions and in three of the seven comparisons the medium difficulty conditions were also enhanced relative to the low difficulty conditions.

Alpha band power changes were driven primarily by the interactions between the difficulty and traffic manipulations at FP1, F7, Fz, F4, and Cz sites with F3 and P4 showing marginally significant interactions ($p < .055$ and $p < .06$). Percent power was highest for the low complexity condition, lower for the high complexity condition and lowest for the medium complexity condition (see Figure 13). Pairwise contrasts revealed significantly higher percent alpha power in the low compared to both the high and medium conditions at sites Fz, F4 and Cz. For the volume manipulations the three workload levels were statistically equivalent. Percent alpha power at T6 also yielded a significant main effect due to task difficulty with greater alpha percent power in the low difficulty conditions compared to both the medium and high difficulty conditions which were not significantly different.

As indicated in Figure 14, the beta 1 band revealed significant differences between the traffic conditions at F3, Fz, F4, Cz, T4 with C3, C4 and P4 marginally significant ($p < .055$ to $p < .066$). The complexity manipulation was associated with a higher percent of Beta 1 power than the volume manipulation. Significant beta 1 interactions were found at sites FP2 and T3. Pairwise contrasts indicated that the overload condition was associated with significantly larger percent beta 1 than with the low volume condition at F7 and T4 or the high volume condition at T6. However, low complexity and medium complexity conditions were associated with greater beta 1 relative power than overload at sites Fz, F3 and Pz and at sites F7 and T4 respectively (Figure 15). For the beta 2 band, there was significantly greater percent power in the complexity traffic than in the volume traffic conditions at site C4 but this band was unaffected by manipulations of workload (Figure 16).

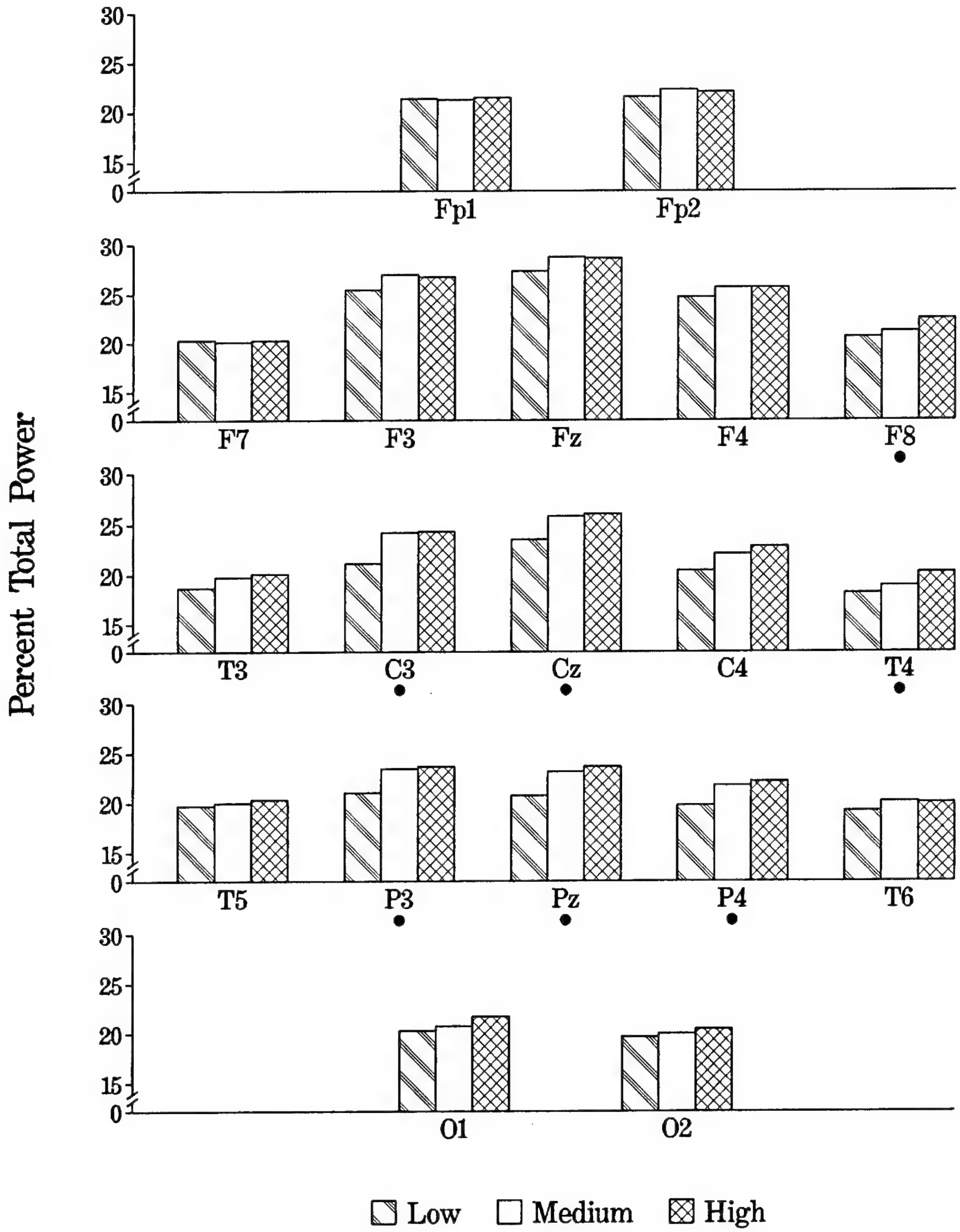



**Figure 12.** Percent total power in the theta band during differing workload levels. Filled circles below and electrode indicate statistical significance at $p < .05$.

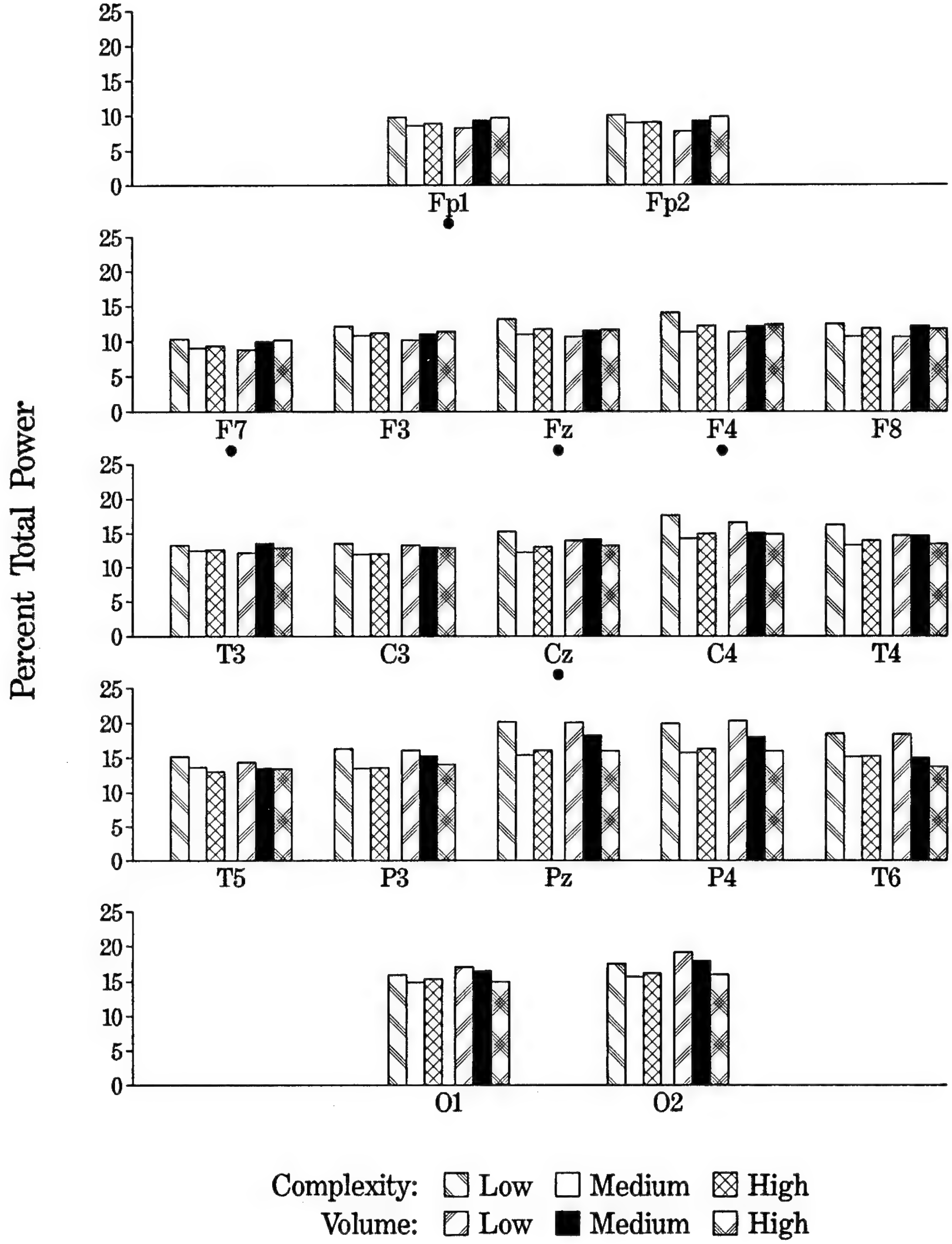



**Figure 13.** Percent total power in the alpha band during all conditions. Filled circles below an electrode indicate significant interactions at $p < .05$.

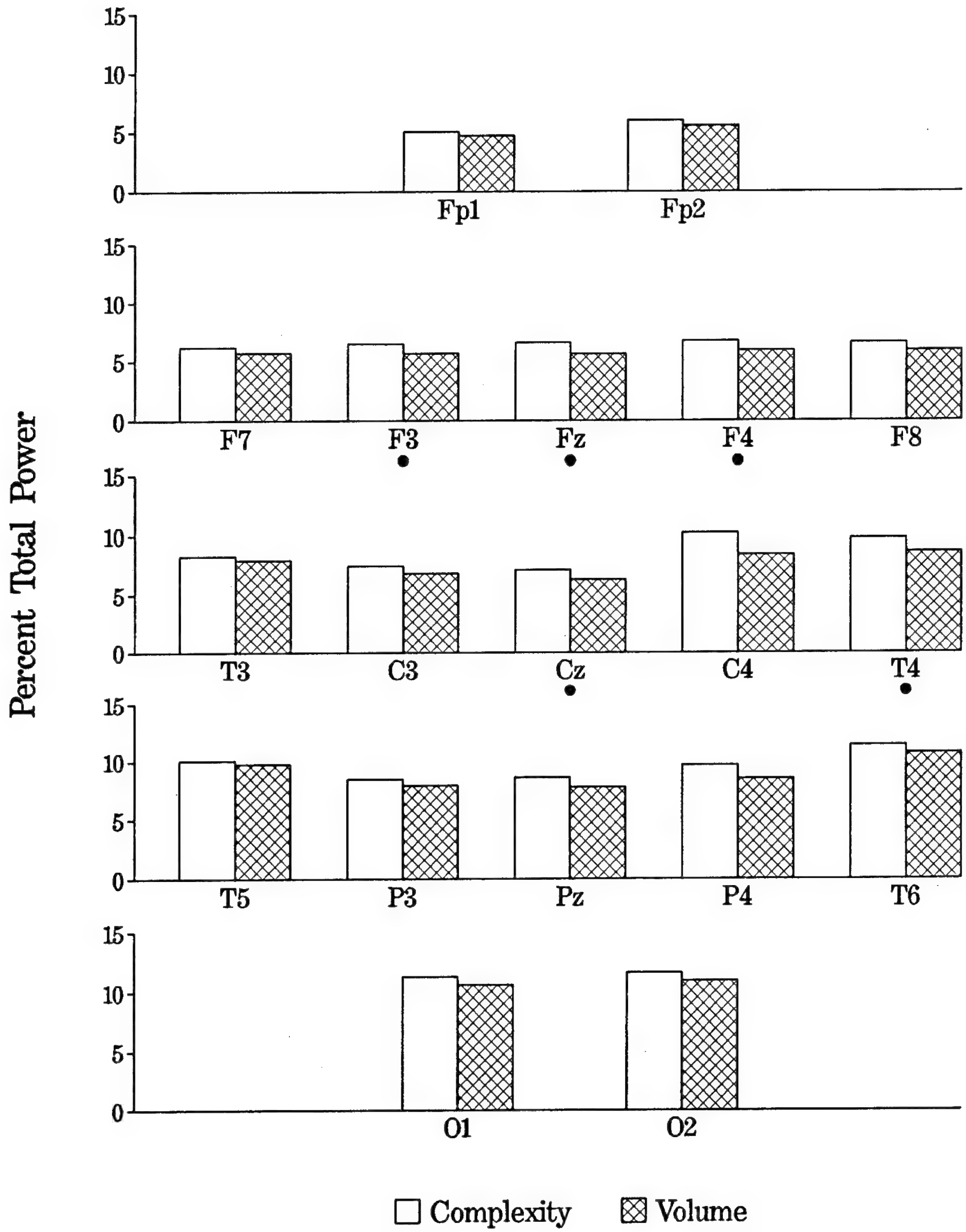



**Figure 14.** Percent total power in the beta 1 band during the volume and complexity scenarios. Filled circles indicate a statistically significant main effect of traffic pattern at $p < .05$.

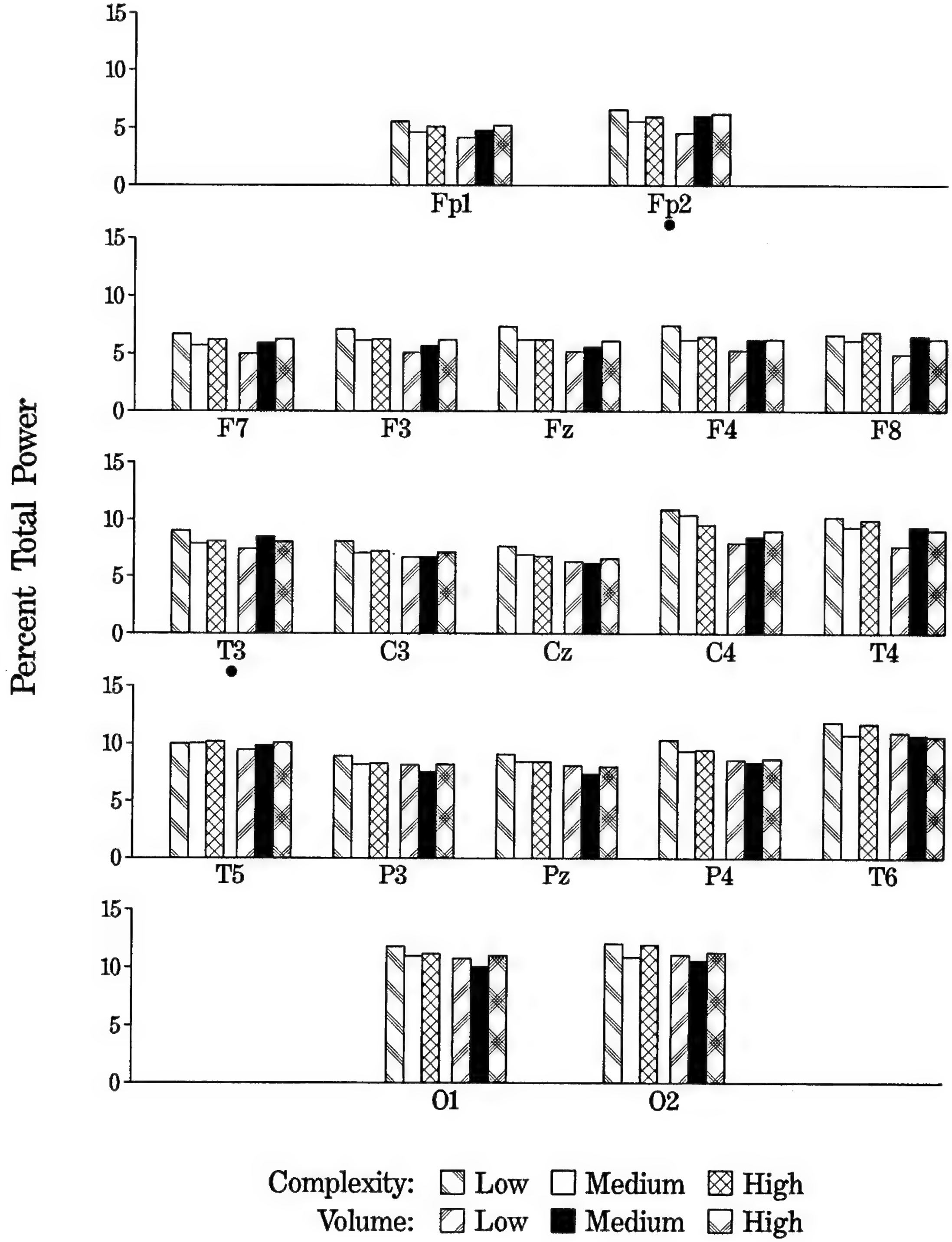



**Figure 15.** Percent total power in the beta 1 band during all conditions. Filled circles below an electrode indicate a statistically significant interaction at $p < .05$.

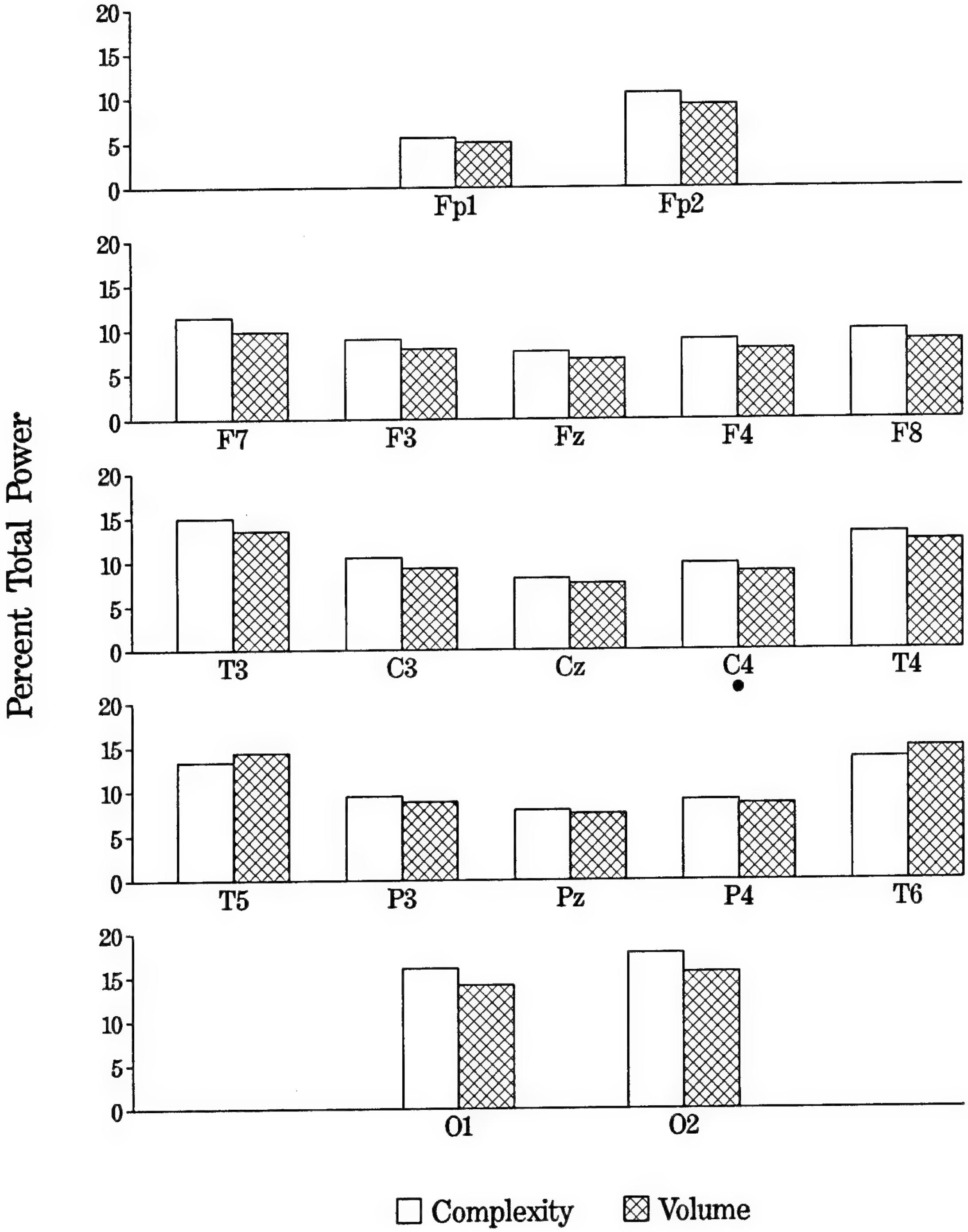



**Figure 16.** Percent total power in the beta 2 band during the traffic volume and complexity scenarios. Significant main effects at the p <.05 level are indicated by filled circles below the electrode.

# DISCUSSION

The results of this study provide support for the sensitivity of a variety of workload indexes to manipulations in the difficulty of a simulated ATC task using Air Force controllers. Specifically, changes in difficulty produced differences in TRACON performance, TLX ratings, eye blink rate, respiration rate, and relative EEG frequency powers. Changes in the traffic manipulation, complexity or volume, produced changes in the EEG only. These differences were consistent with expectations derived from the workload and ATC literature. While the overload condition did represent the highest workload level based on subjective ratings and performance, only one subject reported to have "lost the picture" as anticipated.

The decrease in eye blink rate with higher workload suggests that part of the increased demands placed upon the controllers involves visual attention. The relatively short duration of the flight scenarios should preclude fatigue related increases in blink rates. Furthermore, the lack of significant changes in blink duration and saccades in this study support the premise that fatigue was not a contributing factor since both of these eye measures have been reported to increase concomitantly with time on task and fatigue (Caldwell et al., 1994). This interpretation is also supported by other investigations reporting decreased blink rates when operators were under conditions of increased visual load (Stern & Dunham, 1990; Wilson, et al., 1987; Wilson, 1993). As the visual demands increased, the controllers blinked less often so as not to miss important events on the radar scope.

It is possible that the increased respiration rate associated with increased task demand that was evidenced in this study reflected the increased metabolic demands resulting from the greater mental activity required for task performance. Carroll, Turner and colleagues (Turner & Carroll, 1985; Carroll, Turner & Hellawell, 1986; Carroll, Turner & Rogers, 1987) have reported increases in respiration rates when subjects performed more difficult laboratory tasks. However, they reported correlated changes in heart rate and respiration

rate which were not found in the present study. Backs, Ryan, and Wilson (1994), on the other hand, reported that respiration was more sensitive to changes in tracking task difficulty than was heart rate. Wientjes (1993) also found respiration rate, among other respiration variables, to be more sensitive to changes in the difficulty of a memory loading task than was heart rate and Wientjes (1992) hypothesizes that in tasks such as this changes in respiration rate are due to psychological demands and not metabolic demands. The current data support this interpretation since heart rate did not change during the TRACON manipulations. In more applied environments, Opmeer and Krol (1973) reported respiration rate to be more sensitive to simulated flight workload than is heart rate and Harding (1987) also reported increased respiration rates during the more mentally demanding portions of aircraft flights.

The lack of significant changes in heart rate and heart rate variability suggests that these measures were not sensitive to the variances in mental demand for this particular task. Costa (1993) has also reported that heart rate was not reliably related to ATC workload as measured by the number of aircraft handled during three shifts with varying workload. It appears that in the present study one level of heart rate activity was sufficient to supply the metabolic needs of the brain at all levels of task demand.

The topographically recorded relative EEG power was sensitive to both difficulty and traffic manipulations of the TRACON simulation as well as to the interaction between these manipulations. Furthermore, EEG bands were differentially sensitive to the TRACON difficulty manipulations. In addition, these changes were manifest at different electrode sites across the scalp suggesting that different brain areas were activated.

Significant changes in delta activity were found at frontal and one left temporal electrode site (F3, Fz, F4, T3). The frontal scalp position of these sites supports the proposition that the observed increase in delta power may be due to residual eye activity not eliminated despite the application of an eye blink correction procedure. However, this seems improbable since the pattern of changes does not mirror the pattern of eye blink activity. Eye blink activity reduced as the task became more difficult while the relative delta power

increased during the more difficult segments. In addition, the observed difficulty related patterns of EEG activity were different for the volume and complexity manipulations and were for the most part opposite of that found for the alpha band at the frontal sites. Thus, these data suggest that the relative delta power may be directly influenced by task specific requirements. There is a paucity of discussion of this band in the literature and these results clearly indicate that further examination of the delta band in complex cognitive situations is warranted.

Enhanced theta band activity was seen primarily at central and parietal sites with similar increases apparent at right frontal and right temporal recording sites. Increased mental workload in both of the traffic manipulations was associated with increased theta relative power. The high difficulty situations were always associated with significantly greater relative power than the low difficulty conditions and in three of the seven comparisons the medium difficulty conditions were also enhanced relative to the low difficulty conditions. The relationship of theta activity to human mental state is not clear. Some authors report increased theta activity with increased cognitive activity while others report increased theta when cognitive demands decrease in vigilance situations. Lang and colleagues (Lang, et al., 1987; Lang et al., 1988) have reported increased theta over frontal sites during learning tasks compared to a baseline condition. Mecklinger and colleagues (Mecklinger, Kramer & Strayer, 1992; Pennekamp, et al., 1994) have also reported increased theta in tasks requiring increased attention and memory load while Gevins, et al. (1979) and Ishihara and Yoshi (1972) reported increased theta in several tasks which required memory processing, geometric design evaluation, reading, and spatial processes. Gundel and Wilson (1992) and Pigeau et al. (1987) also found increased theta activity in response to increased task difficulty using several tasks. On the other hand, increased theta has been reported in studies in which subjects become less vigilant due to the repetitive and long term nature of their task. For example, Belyavin and Wright (1987) reported increases in theta activity with increased time on task. While seemingly contradictory, these results may actually indicate that increased theta activity reflects at least two different mechanisms, one which is related to certain classes of cognitive activity, for example, learning or complex task

performance, and another mechanism which is determined by the arousal state of the subject such that decreased arousal and increased fatigue produce increases in theta activity. The topographic distribution of these changes may be useful for determining the underlying cortical areas that are engaged in the task specific increased activity. If this premise is correct, then not only changes in theta but the location of those changes could be predictive of the nature of the cognitive demands experienced in a given situation.

The changes in the alpha band were primarily represented by decreased relative power between the low and medium complexity conditions at sites F7, Fz, F4 and C4 while the cognitive demands of varying levels of traffic volume did not produce changes in alpha. There also was a significant workload-related reduction in alpha from low to medium to high difficulty at the T6 electrode site. Numerous reports have shown decreased alpha with increased task difficulty. Earle & Pikus (1982) reported decreased alpha with increased difficulty of mental arithmetic tasks. Using several laboratory cognitive tasks, Gundel and Wilson (1992) and Pigeau, et al. (1987) found alpha to decrease with increased task difficulty. It is interesting to note that with the TRACON task which involved complex cognitive activity the significant decreases in alpha were between the low and medium difficulty levels and not between low and high.

While the theta band showed significant increases in relative power with increased task difficulty, increased relative beta activity was associated with the complexity manipulations when compared to the volume manipulations primarily over frontal and central sites. This is indicative of increased processing requirements in the complexity condition when several variables had to be dealt with during the simulation. In the volume conditions, the primary task was to manage differing numbers of aircraft while the complexity condition manipulated aircraft type and pilot's ability which may have caused the subjects to utilize different cognitive strategies. Since beta band power appears to be more sensitive to traffic manipulations, this could be used to differentiate the type of processing rather than the level of processing required. There is relatively little in the literature concerning changes to beta due to manipulations of task difficulty. These data suggest that beta levels may be associated

with the nature of the cognitive activity during complex task performance.

The results of this investigation provide support for the sensitivity of a variety of workload indexes to manipulations in the difficulty of a simulated ATC task. Specifically, the workload manipulations produced differences in TRACON performance, TLX ratings, and three physiological measures (eye blink, respiration and brain activity), and the differences were consistent with expectations derived from the workload and ATC literature. The overload condition presented the highest workload but only one subject "lost the picture." The ability of a number of psychophysiological measures to distinguish between ATC task difficulty and traffic manipulations supports the use of these measures to assess controller workload during normal working conditions. The psychophysiological measures contribute information beyond that provided by performance and subjective measures and have the advantage of being continuous and non-intrusive to the controllers' primary responsibilities. In general, these data suggest that, in addition to being more sensitive to task variables than either performance or subjective measures, the EEG are also more sensitive than the other physiological measures used in this study. If true, then EEG recordings could be used to evaluate the relative contributions of workload variables not detected by other metrics. The effects of new procedures and equipment could be assessed by the use of physiological measures in conjunction with performance and subjective measures. Questions concerning the workload of different shifts, different ATC locations, and other related variables could be assessed with the same procedures. The feasibility of utilizing psychophysiological data to correctly classify operator workload in the laboratory (Wilson & Fisher, 1995) and in actual flight (Wilson & Fisher, 1991) has been demonstrated. Systems that can provide on-line reduction of multiple physiological variables and estimate operator workload are currently being developed (Wilson, 1994).

Even though the TRACON task faithfully mirrors several aspects of "real-world" ATC, it does not incorporate all elements of ATC complexity. Despite these limitations, the results of this study support the utility of TRACON as a tool for conducting controlled studies of ATC workload. As Ackerman (1992) observed, the task helps "bridge the gap"

between the simple information-processing tasks used in most laboratory studies and the complex tasks performed by ATCs. In addition, the use of active air traffic controllers as subjects enhances the external validity of these findings. Ackerman (1992) and Ackerman and Kanfer (1993) studied TRACON performance among college students and ATC trainees, respectively, but the current investigation is the first to use TRACON for workload assessment among active ATCs.

## REFERENCES


Ackerman, P. L. (1992). Predicting individual differences in complex skill acquisition: Dynamics of ability determinants. Journal of Applied Psychology, 77, 598-614.

Ackerman, P. L., & Kanfer, R. (1993). Integrating laboratory and field study for improving selection: Development of a battery for predicting air traffic control success. Journal of Applied Psychology, 78, 413-432.

Backs, R. W., Ryan, A. M. & Wilson, G. F. (1994). Psychophysiological measures of workload during continuous manual performance. Human Factors, 36, 514-531.

Bauer, L., Strock, B., Goldstein, R., Stern, J. & Walrath, L. (1985). Auditory discrimination and the eyeblink. Psychophysiology, 22, 636-641.

Belyavin, A. & Wright, N. A. (1987). Changes in electrical activity of the brain with vigilance. Electroencephalography and Clinical Neurophysiology, 66, 137-144.

Caldwell, J.A., Wilson, G.F., Centiguc, M., Galliard, A.W.K., Gundel, A., Lagarde, D., Makeig, S., Myhre, G. & Wright, N.A. (1994). Psychophysiological Assessment Methods, AGARD Advisory Report 324. Paris: AGARD.

Carroll, D., Turner, J. R. & Hellawell, J. C. (1986). Heart rate and oxygen consumption during active psychological challenge: The effects of level of difficulty. Psychophysiology, 23, 174-181.

Carroll, D., Turner, J. R. & Rogers, S. (1987). Heart rate and oxygen consumption during mental arithmetic, a video game, and graded static exercise. Psychophysiology, 24, 112-118.

Casali, J. & Weirwille, W. (1983). A comparison of rating scale, secondary task, physiological, and primary task workload estimation techniques in a simulated flight task emphasizing communications load. Human Factors, 25, 623-642.

Costa, G. (1993). Evaluation of workload in air traffic controllers. Ergonomics, 36, 1111-1120.

Danaher, J. W. (1985). Human error in ATC system operations. Human Factors, 22, 535-545.

Earle, J. B. B., & Pikus, A. (1982). The effect of sex and task difficulty on EEG alpha activity in association with arithmetic. Biological Psychology, 15, 1-14.

Egelund, N. (1982). Spectral analysis of heart rate variability as an indicator of driver fatigue. Ergonomics, 25, 663-672.

Federal Aviation Administration (1985). National air-space system plan: Facilities, equipment and associated development. Washington, DC: FAA.

Gevins, A. S., Zeitlin, G. M., Doyle, J. C., Yingling, C. D., Schaffer, R. E., Callaway, E. & Yeager, C. L. (1979). Electroencephalogram correlates of higher cortical functions. Science, 203, 665-668.

Gevins, A. S., Zeitlin, G. M., Yingling, C. D., Doyle, J. C., Dedon, M. F., Schaffer, R.E.,Roumasset, J. T. & Yeager, C. L. (1979). EEG patterns during 'cognitive' tasks. I. Methodology and analysis of complex behaviors. Electroencephalography and Clinical Neurophysiology, 47, 693-703.

Gratton, G., Coles, M. G. H. & Donchin, E. (1983). A new method for off-line removal of ocular artifact. Electroencephalography and Clinical Neurophysiology, 55, 468-484.

Gundel, A. & Wilson, G. F. (1992). Topographical changes in the ongoing EEG related to the difficulty of mental tasks. Brain Topography, 5, 17-25.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), Human mental workload (pp. 139-183). Amsterdam: North-Holland.

Hopkin, V. D. (1989). Implications of automation on air traffic control. In R. S. Jensen (Ed.), Aviation Psychology (pp. 96-108). Brookfield, VT: Gower.

Ishihara, I. & Yoshi, N. (1972). Multivariate analytic study of EEG and mental activity in juvenile delinquents. Electroencephalography and Clinical Neurophysiology, 33, 71-80.

Itoh, Y., Hayashi, Y., Tskui, I. & Saito, S. (1989). Heart rate variability and subjective mental workload in flight task validity of mental workload measurement using HRV

method, In Smith, M. J. & Salvendy, G. (Eds.), Work with Computers: Organizational, Management, Stress and Health Aspects,209-216. Amsterdam: Elsevier.

Jasper, H.H. (1958). The ten-twenty electrode system of the International Federation. Electroencephalography and Clinical Neurophysiology, 10, 371-375.

Kalsbeek, J. W. H. (1971). Standards of acceptable load in ATC tasks. Ergonomics, 14, 641-650.

Lang, M., Lang, W., Diekmann, V. & Kornhuber, H. H. (1987). The frontal theta rhythm indicating motor and cognitive learning. Current trends in event-related potential research. Electroencephalography and Clinical Neurophysiology Supplement 40, R. Johnson, J. W. Rohrbaugh & R. Parasuraman (Eds). 322-327.

Lang, W., Lang, M., Kornhuber, A., Diekamnn, V. & Kornhuber, H. H. (1988). Event-related EEG-spectra in concept formation task. Human Neurobiology, 6, 295-301.

Mecklinger, A., Kramer, A. F. & Strayer, D. L. (1992). Event related potentials and EEG components in a semantic memory search task. Psychophysiology, 29, 104-119.

Mogford, R. H., Murphy, E. D. & Guttman, J. A. (1994). Using knowledge exploration tools to study airspace complexity in air traffic control. The International Journal of Aviation Psychology, 4, 29-45.

Morrison, R., & Wright, R. H. (1989). ATC control and communications problems: An overview of recent ASRS data. Proceedings of the Fifth International Symposium on Aviation Psychology, 902-907.

Nolan, M. S. (1990). Fundamentals of air traffic control. Belmont, CA: Wadsworth.

Pennekamp, P., Bosel, R, Mecklinger, A. & Ott, H. (1994). Differences in EEG-theta for responded and omitted targets in a sustained attention task. Journal of Psychophysiology, 131-141.

Pigeau, R. A., Hoffmann, R., Purcell, S. & Moffitt, A. (1988). The effect of endogenous alpha on hemispheric asymmetries and the relationship of frontal theta to sustained attention. In: Electric and Magnetic Activity of the Central Nervous System: Research and Clinical Applications in Aerospace Medicine, AGARD-CP-432, 20/1-20/16, Neuilly sur Sein, France: NATO.

Redding, R. E. (1992). Analysis of operational errors and workload in air traffic control.Proceedings of the 36th Annual Meeting of the Human Factors Society, 1321-1325.

Roscoe, A. H. (1987). In-flight assessment of workload using pilot ratings and heart rate. In Roscoe, A. H. (Ed.), The Practical Assessment of Pilot Workload, AGARDograph No. 282, 1-14. Neuilly sur Seine, France: AGARD.

Speyer, J. J., Fort, A., Fouillot, J, & Blomberg, R. D. (1988). Dynamic methods for assessing workload for minimum crew certification. In Roscoe, A. H. & Muir, H. C. (Eds.), Workload in Transport Operations, Report number 1B 316-88-06. 196-220. Cologne: DFVLR.

Stager, P., Hameluck, D. & Jubis, R. (1989). Underlying factors in air traffic control incidents. Proceedings of the 33rd Annual Meeting of the Human Factors Society, 43-46.

Stern, J. A. & Dunham, D. N. (1990). The ocular system. In: Cacioppo, J. T., & Tassinary, L. G. (Eds.) Principles of Psychophysiology: Physical, Social and Inferential Elements, 513-553. Cambridge: Cambridge University Press.

Stern, J.A. & Skelly, J. (1984). The eyeblink and workload considerations. Proceedings of the 28th Annual Meeting of the Human Factors Society, 942-944.

Turner, J.R. & Carroll, D. (1985). Heart rate and oxygen consumption during mental arithmetic, a video game, and graded exercise: Further evidence of metabolically-exaggerated cardiac adjustments? Psychophysiology, 22, 261-267.

Ullsperger, P. (Ed). (1993). Psychophysiology of Mental Workload, Schriftenreihe der Bundersanstalt fur Arbeitmedizin, Verlag: Berlin.

Wierwille, W., Rahimi, M & Casali, J. (1985). Evaluation of 16 measures of mental workload using a simulated flight task emphasizing mediational activity. Human Factors, 28, 489-502.

Wientjes, C. J. E. (1992). Respiration in psychophysiology: Methods and applications. Biological Psychology, 34, 179-204.

Wientjes, C. J. E. (1993). Respiration and Stress. Doctoral dissertation, Tilburg: University of Tilburg.

Wilson, G. F. (1992). Applied use of cardiac and respiration measures: Practical considerations and precautions. Biological Psychology, 34, 163-178.

Wilson, G. F. (1993). Air-to-ground training missions: A psychophysiological workload analysis. Ergonomics, 36, 1071-1087.

Wilson, G. F. (1994). Workload assessment monitor (WAM). Proceedings of the Human Factors and Ergonomics Society, 944.

Wilson, G. F. & Eggemeier, F. T. (1991). Psychophysiological assessment of workload in multi-task environments. In Damos, D. L. (Ed.), Multiple-task Performance, 229-260, London: Taylor & Francis.

Wilson, G. F. & Fisher, F. (1991). The use of cardiac and eye blink measures to determine flight segment in F4 crews. Aviation, Space, and Environmental Medicine, 62, 959-962.

Wilson, G. F. & Fisher, F. (1995). Cognitive Task Classification Based Upon Topographic EEG Data. Biological Psychology, 40, 239-250.

Wilson, G. F., & Oliver, C. G. (1991). PATS: Psychophysiological Assessment Test System. In E. Farmer (Ed.), Stress and error in aviation: Proceedings of the XVIII WEAAP Conference (Vol. 2, pp. 15-25). Worcester: Billing & Son Ltd.

Wilson, G. F., Purvis, B., Skelly, J., Fullenkamp, P., & Davis, I. (1987). Physiological data used to measure pilot workload in actual flight and simulator conditions. Proceedings of the 31st Annual Meeting of the Human Factors Society, 779-783..

U. S. Government Printing Office 1996 750-071/00182